



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 09:28 PM BST

PDB ID : 2JTC
Title : 3D structure and backbone dynamics of SPE B
Authors : Chuang, W.; Wang, C.; Houn, H.; Chen, C.; Wang, P.
Deposited on : 2007-07-26

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

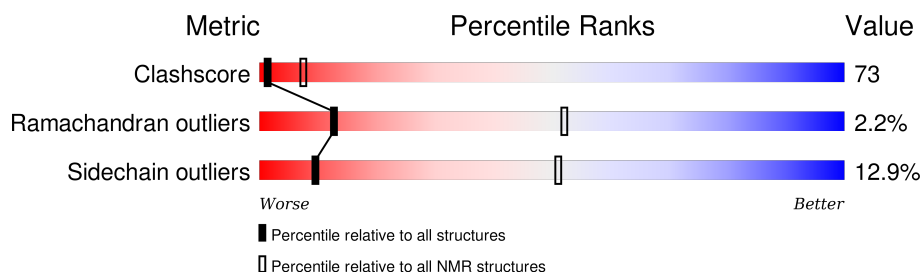
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	253	

2 Ensemble composition and analysis

This entry contains 20 models. Model 8 is the overall representative, medoid model (most similar to other models). The authors have identified model 5 as representative, based on the following criterion: *closest to the average*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:4-A:31, A:37-A:212, A:245-A:253 (213)	0.21	8

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 1 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20
Single-model clusters	4; 13

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3782 atoms, of which 1833 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Streptopain.

Mol	Chain	Residues	Atoms						Trace
1	A	253	Total	C	H	N	O	S	0
			3782	1227	1833	343	375	4	

There is a discrepancy between the modelled and reference sequences:

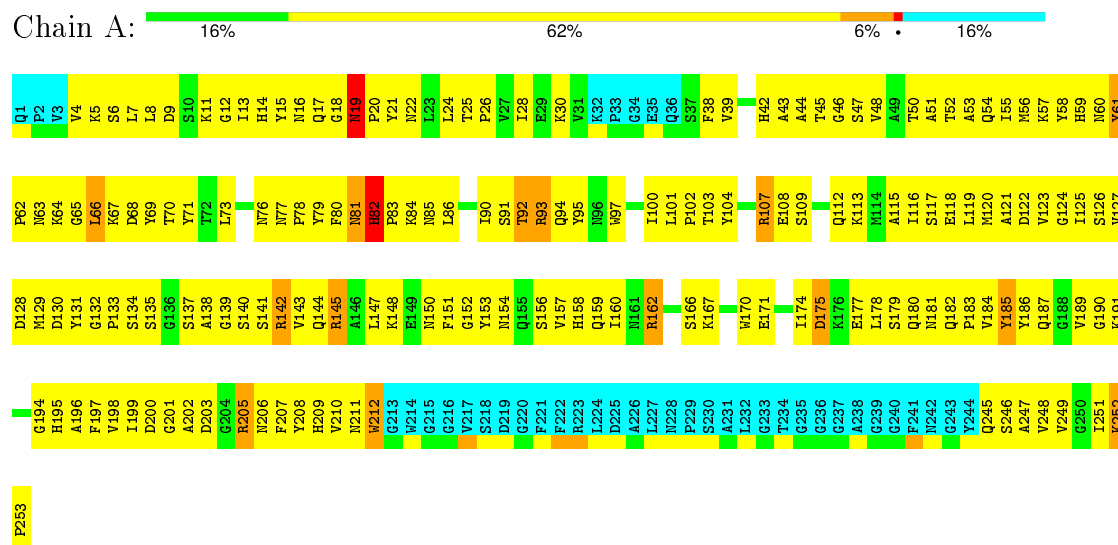
Chain	Residue	Modelled	Actual	Comment	Reference
A	47	SER	CYS	ENGINEERED	UNP P0C0J1

4 Residue-property plots [i](#)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

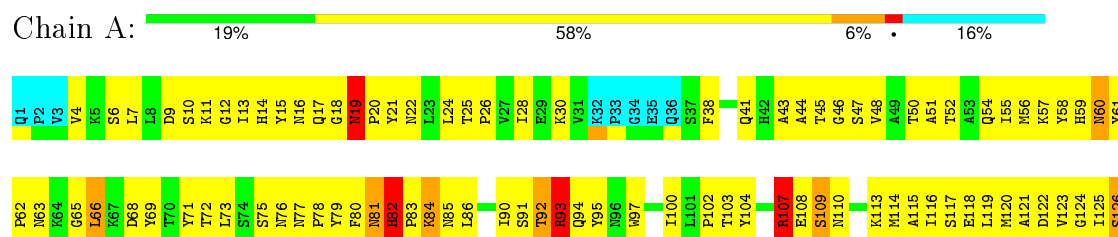
- Molecule 1: Streptopain



4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 8. Colouring as in section 4.1 above.

- Molecule 1: Streptopain



V127	G190		
D128	R191		
M129	V192		
D130	G193		
Y131	G194		
G132	H195		
P133	A196		
S134	F197		
S135	V198		
G136	I199		
S137	D200		
A138	G201		
G139	A202		
S140	D203		
S141	G204		
R142	R205		
V143	N206		
Q144	F207		
R145	Y208		
A146	H209		
L147	V210		
K148	N211		
	W212		
F151	G213		
G152	W214		
Y153	G215		
N154	G216		
Q155	V217		
S156	S218		
V157	D219		
H158	G220		
Q159	F221		
I160	F222		
N161	R223		
R162	L224		
	D225		
S166	A226		
K167	L227		
Q168	N228		
D169	F229		
W170	S230		
E171	A231		
A172	L232		
Q173	G233		
I174	T234		
D175	G235		
K176	G236		
E177	G237		
L178	A238		
S179	G239		
Q180	G240		
N181	F241		
Q182	N242		
P183	G243		
V184	Y244		
Y185	Q245		
Y186	S246		
Q187	A247		
G188	V248		
V189	V249		

G250
I251
K252
P253

5 Refinement protocol and experimental data overview

The models were refined using the following method: *DGSA-distance geometry simulated annealing*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR	refinement	3.185

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality

6.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.98±0.00	0±0/1712 (0.0±0.0%)	0.96±0.01	0±0/2321 (0.0±0.0%)
All	All	0.98	0/34240 (0.0%)	0.96	4/46420 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	5.5±0.7
All	All	0	110

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	69	TYR	CB-CG-CD1	-6.45	117.13	121.00	1	1
1	A	38	PHE	CB-CG-CD2	-5.68	116.83	120.80	16	3

There are no chirality outliers.

5 of 6 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	205	ARG	Sidechain	20
1	A	145	ARG	Sidechain	20
1	A	107	ARG	Sidechain	20
1	A	142	ARG	Sidechain	19
1	A	162	ARG	Sidechain	17

6.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1670	1580	1599	240±13
All	All	33400	31600	31980	4801

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 73.

5 of 944 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:55:ILE:HG21	1:A:251:ILE:HD13	1.04	1.28	16	4
1:A:116:ILE:HD11	1:A:120:MET:HE2	0.98	1.34	18	2
1:A:160:ILE:HD12	1:A:170:TRP:CH2	0.97	1.94	9	6
1:A:55:ILE:HD11	1:A:198:VAL:HG22	0.96	1.34	4	18
1:A:174:ILE:HG23	1:A:184:VAL:HG21	0.93	1.41	8	14

6.3 Torsion angles [i](#)

6.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	212/253 (84%)	184±3 (87±1%)	24±2 (11±1%)	5±1 (2±0%)	13	52
All	All	4240/5060 (84%)	3673 (87%)	473 (11%)	94 (2%)	13	52

5 of 7 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	81	ASN	20
1	A	19	ASN	20
1	A	82	HIS	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	212	TRP	18
1	A	61	TYR	10

6.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	179/204 (88%)	156±4 (87±2%)	23±4 (13±2%)	9	51
All	All	3580/4080 (88%)	3119 (87%)	461 (13%)	9	51

5 of 79 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	19	ASN	20
1	A	175	ASP	19
1	A	82	HIS	19
1	A	66	LEU	18
1	A	92	THR	18

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided